Information processing in gene regulatory cascades

Aleksandra M. Walczak[1], Andrew Mugler[2], and Chris H. Wiggins[3]

***Short Abstract* — We study the information processing abilities of noisy gene regulatory cascades by optimizing the mutual information between the input and output protein concentrations of systems of various lengths over the parameters of the regulatory functions and input distribution. We find, for threshold regulation, that a cascade of strong regulations converts a unimodal input to a bimodal output and that multimodal inputs are no more informative than bimodal inputs. The capacity to transmit information of up-regulating and down-regulating cascades is the same. However down-regulation is a more costly form of regulation, in terms of the produced number of protein copies.**

***Keywords* — stochastic gene regulation, signaling cascades, information processing**

I. Purpose

Genetic signaling cascades regulate many cellular processes [1] from development [2] to quorum sensing [3]. Downstream genes respond to concentrations of transcription factors produced by upstream genes. Such many gene systems transmit signals by progressive accumulation of transcription factors, translating the concentration of the input protein into the concentration of the output protein. Gene regulatory cascades are composed of genes and proteins, both of which are present in small numbers in the cell. The noise due to small molecule counts adds to the randomness of chemical reactions. Probabilistic approaches have proven necessary to fully account for the variability of molecule numbers within a homogenous population of cells.

We use mutual information between the input and output protein concentrations as a quantitative measure to study the ability of gene regulatory cascades to process signals. We present a study of the properties of the most informative noisy cascades of various lengths.

II. The Model

We describe a cascade using a joint probability distribution of the copy count of proteins produced by each gene in the system. Within the model, we use a birth-death process to describe protein synthesis and degradation. The synthesis rate of the downstream gene depends on the copy count of the proteins produced by the gene directly upstream in the cascade. For concreteness, we focus on the specific case of discontinuous threshold regulation: the downstream species is created at a high or low rate depending on the concentration of the upstream gene.

We find the steady state joint probability distributions of the cascade using an eigenfunction decomposition [4], within the approximation that the behaviour of upstream nodes depends weakly on downstream nodes. We exploit the computational efficiency of this method of evaluating the steady state distributions, to optimize the mutual information between the input and output of cascades, with Poisson input distributions, over the parameters of the model.

III. Results

We find that in the case of threshold regulation with a unimodal input, both unimodal and bimodal output distributions can be the most informative. Both in the case of cascades where the proteins act as repressors (down-regulation) and as activators (up-regulation), for large difference in the synthesis rates between the low and high states, the bimodal output is most informative, whereas for small difference in the synthesis rates between the low and high states the output distribution is unimodal. Larger discontinuities in the production rates guarantee stable bimodality of the input for shorter cascades. Curiously, the capacity of a cascade to transmit information is the same for up-regulating and down-regulating cascades, and depends only on the value of the discontinuity between the production rates. However the production cost for the proteins in the system is different: down-regulating cascades produce more proteins to achieve the optimal capacity compared to up-regulating cascades. We find that, in the case of threshold regulation, the most optimal input for a cascade of any length is always bimodal; a multimodal input offers no further processing power.

[1]Princeton Center for Theoretical Physics, Princeton University, NJ 08540. E-mail: awalczak@princeton.edu

[2]Department of Physics, Columbia University, New York, NY 10027 USA.

[3]Department of Applied Physics and Mathematics, Columbia University, New York, NY 10027 USA.

References

[1] Gomperts, BD, Kramer, IM, Tatham P.E.R, (2002) *Signal Transduction*, Academic Press, San Diego

[2] H. Bolouri, Davidson EH, (2003) Transcriptional regulatory cascades in development: Initial rates, not steady state, determine network kinetics *PNAS* **100**, 9371.

[3] Bassler, BL, (1999) How bacteria talk to each other: regulation of gene expression by quorum sensing *Current Opinion in Microbiology* **2, 582**

[4] Walczak, AM, Mugler, A, Wiggins, CH, (2008) A stochastic spectral analysis of transcriptional regulatory cascades, *PNAS* 2008-11999R.